



OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/898,216

DATE: 01/25/2002

TIME: 16:58:04

Input Set : N:\Crif3\RULE60\09898216.raw

Output Set : N:\CRF3\01252002\I898216.raw

## SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Hillman, Jennifer L.

6 Goli, Surya K.

8 (ii) TITLE OF INVENTION: NOVEL HUMAN MEMBRANE PROTEIN

10 (iii) NUMBER OF SEQUENCES: 7

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

14 (B) STREET: 3174 Porter Drive

15 (C) CITY: Palo Alto

16 (D) STATE: CA

17 (E) COUNTRY: USA

18 (F) ZIP: 94304

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Diskette

22 (B) COMPUTER: IBM Compatible

23 (C) OPERATING SYSTEM: DOS

24 (D) SOFTWARE: FastSEQ for Windows Version 2.0

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/09/898,216

C--> 28 (B) FILING DATE: 02-Jul-2001

29 (C) CLASSIFICATION:

31 (vii) PRIOR APPLICATION DATA:

32 (A) APPLICATION NUMBER: 08/781,562

33 (B) FILING DATE:

35 (viii) ATTORNEY/AGENT INFORMATION:

36 (A) NAME: Billings, Lucy J.

37 (B) REGISTRATION NUMBER: 36,749

38 (C) REFERENCE/DOCKET NUMBER: PF-0181 US

40 (ix) TELECOMMUNICATION INFORMATION:

41 (A) TELEPHONE: 415-855-0555

42 (B) TELEFAX: 415-845-4166

43 (C) TELEX:

46 (2) INFORMATION FOR SEQ ID NO: 1:

48 (i) SEQUENCE CHARACTERISTICS:

49 (A) LENGTH: 356 amino acids

50 (B) TYPE: amino acid

51 (C) STRANDEDNESS: single

52 (D) TOPOLOGY: linear

54 (vii) IMMEDIATE SOURCE:

55 (A) LIBRARY: Consensus

56 (B) CLONE: Consensus

58 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ENTERED

## RAW SEQUENCE LISTING

DATE: 01/25/2002

PATENT APPLICATION: US/09/898,216

TIME: 16:58:04

Input Set : N:\Crif3\RULE60\09898216.raw

Output Set: N:\CRF3\01252002\I898216.raw

```

60 Met Leu Ala Arg Ala Ala Arg Gly His Trp Gly Pro Phe Ala Glu Gly
61 1 5 10 15
62 Leu Ser Thr Gly Phe Trp Pro Arg Ser Gly Arg Ala Ser Ser Gly Leu
63 20 25 30
64 Pro Arg Asn Thr Val Val Leu Phe Val Pro Gln Gln Glu Ala Trp Val
65 35 40 45
66 Val Glu Arg Met Gly Arg Phe His Arg Ile Leu Glu Pro Gly Leu Asn
67 50 55 60
68 Ile Leu Ile Pro Val Leu Asp Arg Ile Arg Tyr Val Gln Ser Leu Lys
69 65 70 75 80
70 Glu Ile Val Ile Asn Val Pro Glu Gln Ser Ala Val Thr Leu Asp Asn
71 85 90 95
72 Val Thr Leu Gln Ile Asp Gly Val Leu Tyr Leu Arg Ile Met Asp Pro
73 100 105 110
74 Tyr Lys Ala Ser Tyr Gly Val Glu Asp Pro Glu Tyr Ala Val Thr Gln
75 115 120 125
W--> 76 Leu Ala Gln Thr Thr Met Arg Ser Glu Leu Gly Lys Leu Ser Xaa Asp
77 130 135 140
78 Lys Val Phe Arg Glu Arg Glu Ser Leu Asn Ala Ser Ile Val Asp Ala
79 145 150 155 160
80 Ile Asn Gln Ala Ala Asp Cys Trp Gly Ile Arg Cys Leu Arg Tyr Glu
81 165 170 175
82 Ile Lys Asp Ile His Val Pro Pro Arg Val Lys Glu Ser Met Gln Met
83 180 185 190
84 Gln Val Glu Ala Glu Arg Arg Lys Arg Ala Thr Val Leu Glu Ser Glu
85 195 200 205
86 Gly Thr Arg Glu Ser Ala Ile Asn Val Ala Glu Gly Lys Lys Gln Ala
87 210 215 220
88 Gln Ile Leu Ala Ser Glu Ala Glu Lys Ala Glu Gln Ile Asn Gln Ala
89 225 230 235 240
90 Ala Gly Glu Ala Ser Ala Val Leu Ala Lys Ala Lys Ala Lys Ala Glu
91 245 250 255
92 Ala Ile Arg Ile Leu Ala Ala Ala Leu Thr Gln His Asn Gly Asp Ala
93 260 265 270
94 Ala Ala Ser Leu Thr Val Ala Glu Tyr Val Ser Ala Phe Ser Lys
95 275 280 285
96 Leu Ala Lys Asp Ser Asn Thr Ile Leu Leu Pro Ser Asn Pro Gly Asp
97 290 295 300
98 Val Thr Ser Met Val Ala Gln Ala Met Gly Val Tyr Gly Ala Leu Thr
99 305 310 315 320
100 Lys Ala Pro Val Pro Gly Thr Pro Asp Ser Leu Ser Ser Gly Ser Ser
101 325 330 335
W--> 102 Arg Asp Val Gln Gly Thr Asp Ala Ser Xaa Asp Glu Glu Leu Asp Arg
103 340 345 350
104 Val Lys Met Ser
105 355
107 (2) INFORMATION FOR SEQ ID NO: 2:
109 (i) SEQUENCE CHARACTERISTICS:
110 (A) LENGTH: 1188 base pairs

```

## RAW SEQUENCE LISTING

DATE: 01/25/2002

PATENT APPLICATION: US/09/898,216

TIME: 16:58:04

Input Set : N:\Crf3\RULE60\09898216.raw

Output Set: N:\CRF3\01252002\I898216.raw

```

111      (B) TYPE: nucleic acid
112      (C) STRANDEDNESS: single
113      (D) TOPOLOGY: linear
115      (vii) IMMEDIATE SOURCE:
116          (A) LIBRARY: Consensus
117          (B) CLONE: Consensus
119      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
121  GGCTTCTGGG AGCNACCGCT CCGCTCGTCT CGTTGGTTCC GGAGGTCGCT GCGGCGGTGG      60
122  GAAATGCTGG CGCGCGCGGC GCGGGGGCAC TGGGGCCCTT TTGCTGAGGG GCTCTCTACT      120
123  GGCTTCTGGC CGCGCTCCGG CCGCGCCTCC TCTGGATTGC CCCGAAACAC CGTGGTACTG      180
124  TTCGTGCCGC AGCAGGAGGC CTGGGTGGTG GAGCGAATGG GCCGATTCCA CCGGATCCTG      240
125  GAGCCTGGTT TGAACATCCT CATCCCTGTG TTAGACCGGA TCCGATATGT GCAGAGTCTC      300
126  AAGGAAATTG TCAATCAACGT GCCTGAGCAG TCGGCTGTGA CTCTCGACAA TGTAACCTCTG      360
127  CAAATCGATG GAGTCCTTTA CCTGCGCATC ATGGACCCTT ACAAGGCAAG CTACGGTGTG      420
128  GAGGACCCTG AGTATGCCGT CACCCAGCTA GCTCAAACAA CCATGAGATC AGAGCTCGGC      480
129  AAATCTCTN TGGACAAAGT CTTCCGGGAA CGGGAGTCCC TGAATGCCAG CATTGTGGAT      540
130  GCCATCAACC AAGCTGCTGA CTGCTGGGGT ATCCGCTGCC TNCGTTATGA GATCAAGGAT      600
131  ATCCATGTGC CACCCCGGGT GAAAGAGTCT ATGCAGATGC AGGTGGAGGC AGAGCGGCGG      660
132  AAACGGGCCA CAGTTCTAGA GTCTGAGGGG ACCCGAGAGT CGGCCATCAA TGTGGCAGAA      720
133  GGAAGAAGAC AGGCCAGAT CTTGGCCTCC GAAGCAGAAA AGGCTGAACA GATAAATCAG      780
134  GCAGCAGGAG AGGCCAGTGC AGTTCTGGCG AAGGCCAAGG CTAAAGCTGA AGCTATTCTGA      840
135  ATCCTGGCTG CAGCTCTGAC ACAACATAAT GGAGATGCAG CAGCTTCACT GACTGTGGCC      900
136  GAGCAGTATG TCAGCGCGTT CTCCAAACTG GCCAAGGACT CCAACACTAT CCTACTGCCC      960
137  TCCAACCCTG GCGATGTCAC CAGCATGGTG GCTCAGGCCA TGGGTGTATA TGGAGCCCTC     1020
138  ACCAAAGCCC CAGTGCCAGG GACTCCAGAC TCACTCTCCA GTGGGAGCAG CAGAGATGTC     1080
139  CAGGGTACAG ATGCAAGTNT TGATGAGGAA CTTGATCGAG TCAAGATGAG TTAGTGGAGC     1140
140  TGGGCTTNGC CAGGGAGTCT GGGGACAAGG AAGCAGATTT TCCTGATT      1188
142      (2) INFORMATION FOR SEQ ID NO: 3:
144          (i) SEQUENCE CHARACTERISTICS:
145              (A) LENGTH: 288 amino acids
146              (B) TYPE: amino acid
147              (C) STRANDEDNESS: single
148              (D) TOPOLOGY: linear
150          (vii) IMMEDIATE SOURCE:
151              (A) LIBRARY: Genbank
152              (B) CLONE: 31069
154          (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
156  Met Ala Glu Lys Arg His Thr Arg Asp Ser Glu Ala Gln Arg Leu Pro
157   1          5          10          15
158  Asp Ser Phe Lys Asp Ser Pro Ser Lys Gly Leu Gly Pro Cys Gly Trp
159   20          25          30
160  Ile Leu Val Ala Phe Ser Phe Leu Phe Thr Val Ile Thr Phe Pro Ile
161   35          40          45
162  Ser Ile Trp Met Cys Ile Lys Ile Ile Lys Glu Tyr Glu Arg Ala Ile
163   50          55          60
164  Ile Phe Arg Leu Gly Arg Ile Leu Gln Gly Gly Ala Lys Gly Pro Gly
165   65          70          75          80
166  Leu Phe Phe Ile Leu Pro Cys Thr Asp Ser Phe Ile Lys Val Asp Met
167   85          90          95

```

## RAW SEQUENCE LISTING

DATE: 01/25/2002

PATENT APPLICATION: US/09/898,216

TIME: 16:58:04

Input Set : N:\Crf3\RULE60\09898216.raw

Output Set: N:\CRF3\01252002\I898216.raw

```

168 Arg Thr Ile Ser Phe Asp Ile Pro Pro Gln Glu Ile Leu Thr Lys Asp
169           100           105           110
170 Ser Val Thr Ile Ser Val Asp Gly Val Val Tyr Tyr Arg Val Gln Asn
171           115           120           125
172 Ala Thr Leu Ala Val Ala Asn Ile Thr Asn Ala Asp Ser Ala Thr Arg
173           130           135           140
174 Leu Leu Ala Gln Thr Thr Leu Arg Asn Val Leu Gly Thr Lys Asn Leu
175           145           150           155           160
176 Ser Gln Ile Leu Ser Asp Arg Glu Glu Ile Ala His Asn Met Gln Ser
177           165           170           175
178 Thr Leu Asp Asp Ala Thr Asp Ala Trp Gly Ile Lys Val Glu Arg Val
179           180           185           190
180 Glu Ile Lys Asp Val Lys Leu Pro Val Gln Leu Gln Arg Ala Met Ala
181           195           200           205
182 Ala Glu Ala Glu Ala Ser Arg Glu Ala Arg Ala Lys Val Ile Ala Ala
183           210           215           220
184 Glu Gly Glu Met Asn Ala Ser Arg Ala Leu Lys Glu Ala Ser Met Val
185           225           230           235           240
186 Ile Thr Glu Ser Pro Ala Ala Leu Gln Leu Arg Tyr Leu Gln Thr Leu
187           245           250           255
188 Thr Thr Ile Ala Ala Glu Lys Asn Ser Thr Ile Val Phe Pro Leu Pro
189           260           265           270
190 Ile Asp Met Leu Gln Gly Ile Ile Gly Ala Lys His Ser His Leu Gly
191           275           280           285

```

## 193 (2) INFORMATION FOR SEQ ID NO: 4:

## 195 (i) SEQUENCE CHARACTERISTICS:

196 (A) LENGTH: 280 amino acids

197 (B) TYPE: amino acid

198 (C) STRANDEDNESS: single

199 (D) TOPOLOGY: linear

## 201 (vii) IMMEDIATE SOURCE:

202 (A) LIBRARY: GenBank

203 (B) CLONE: 1065452

## 205 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```

207 Met Asn Leu Lys Thr Cys Ser Leu Ser Thr His Ser Phe Leu Gln Lys
208 1           5           10           15
209 Lys Asn Glu Lys His Asp Gly Asn Pro Glu His Tyr Asp Thr Gly Leu
210           20           25           30
211 Gly Phe Cys Gly Trp Phe Leu Met Gly Leu Ser Trp Ile Met Val Ile
212           35           40           45
213 Ser Thr Phe Pro Val Ser Ile Tyr Phe Cys Met Lys Val Val Gln Glu
214           50           55           60
215 Tyr Glu Arg Ala Val Ile Phe Arg Leu Gly Arg Leu Ile Gly Gly Gly
216           65           70           75           80
217 Ala Lys Gly Pro Gly Ile Phe Phe Val Leu Pro Cys Ile Glu Ser Tyr
218           85           90           95
219 Thr Lys Val Asp Leu Arg Thr Val Ser Phe Ser Val Pro Pro Gln Glu
220           100          105          110
221 Ile Leu Thr Lys Asp Ser Val Thr Thr Ser Val Asp Ala Val Ile Tyr

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/898,216

DATE: 01/25/2002

TIME: 16:58:04

Input Set : N:\Crif3\RULE60\09898216.raw

Output Set: N:\CRF3\01252002\I898216.raw

```

222          115          120          125
223 Tyr Arg Ile Ser Asn Ala Thr Val Ser Val Ala Asn Val Glu Asn Ala
224          130          135          140
225 His His Ser Thr Arg Leu Leu Ala Gln Thr Thr Leu Arg Asn Met Leu
226          145          150          155          160
227 Gly Thr Arg Ser Leu Ser Glu Ile Leu Ser Asp Arg Glu Thr Leu Ala
228          165          170          175
229 Ala Ser Met Gln Thr Ile Leu Asp Glu Ala Thr Glu Ser Trp Gly Ile
230          180          185          190
231 Lys Val Glu Arg Val Glu Ile Lys Asp Val Arg Leu Pro Ile Gln Leu
232          195          200          205
233 Gln Arg Ala Met Ala Ala Glu Ala Glu Ala Thr Arg Glu Ala Arg Ala
234          210          215          220
235 Lys Val Ile Ala Ala Glu Gly Glu Gln Lys Ala Ser Arg Ala Leu Arg
236          225          230          235          240
237 Asp Ala Ala Ser Val Ile Ala Gln Ser Pro Ala Ala Leu Gln Leu Arg
238          245          250          255
239 Tyr Leu Gln Thr Leu Asn Ser Val Ala Arg Glu Lys Phe Asp Asp His
240          260          265          270
241 Leu Pro Thr Ser Asp Gly Ile Ser
242          275          280
244 (2) INFORMATION FOR SEQ ID NO: 5:
246 (i) SEQUENCE CHARACTERISTICS:
247 (A) LENGTH: 415 amino acids
248 (B) TYPE: amino acid
249 (C) STRANDEDNESS: single
250 (D) TOPOLOGY: linear
252 (vii) IMMEDIATE SOURCE:
253 (A) LIBRARY: GenBank
254 (B) CLONE: 1353669
256 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
258 Met Glu Tyr Gly Met Pro Glu Gly Ser Tyr Asp Ser Val Phe Thr Tyr
259 1 5 10 15
260 Ala Pro Tyr Asn Asp Leu Asp Lys Met Gly Tyr Met Gly Pro Ala Arg
261 20 25 30
262 Gln Gly Met Met Leu Gly Asn Lys Tyr Gly Asn Phe Thr Tyr Thr Arg
263 35 40 45
264 Asp Tyr Gly Val Asn Met Glu Asp Asp Ile Lys Pro Leu Ser Ala Ile
265 50 55 60
266 Glu Leu Leu Ile Phe Cys Val Ser Phe Leu Phe Val Val Met Thr Met
267 65 70 75 80
268 Pro Leu Ser Leu Leu Phe Ala Leu Lys Phe Ile Ser Thr Ser Glu Lys
269 85 90 95
270 Leu Val Val Leu Arg Leu Gly Arg Ala Gln Lys Thr Arg Gly Pro Gly
271 100 105 110
272 Ile Thr Leu Val Ile Pro Cys Ile Asp Thr Thr His Lys Val Thr Met
273 115 120 125
274 Ser Ile Thr Ala Phe Asn Val Pro Pro Leu Gln Ile Ile Thr Thr Asp
275 130 135 140

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/898,216

DATE: 01/25/2002

TIME: 16:58:05

Input Set : N:\Crf3\RULE60\09898216.raw

Output Set: N:\CRF3\01252002\I898216.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:76 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1

L:102 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1